Oxford University Press
London Edinburgh Glasgow Copenhagen
New York Toronto Melbourne Cape Town
Bombay Calcutta Madras Shanghai
Humphrey Milford Publisher to the UNIVERSITY

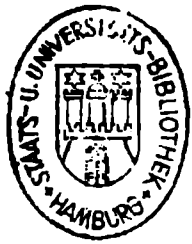
THE CALCULUS OF VARIANTS

An Essay on Textual Criticism

By

W. W. GREG

OXFORD
AT THE CLARENDON PRESS
1927



It is not going too far to say that the announcement that physicists would have in future to study the theory of tensors created a veritable panic among them when the verification of Einstein's predictions was first announced.—A. N. WHITEHEAD.

PREFACE

THE subject considered in the following pages, under the rather pretentious title of the Calculus of Variants, has been the central problem of textual criticism at any rate since the establishment of the genealogical method. I am not here concerned to inquire whether that problem is completely soluble, though I have been unable to avoid the question altogether, but only to suggest the use of more rigorous and in the end simpler methods of approach. A considerable gain in ease and certainty can, I believe, be attained by a partial substitution of formal rules for the continuous application of reason; and I have been driven to seek it because in practice I always myself feel considerable uncertainty as to what can and what cannot be legitimately inferred from a particular set of variants, and observation leads me to doubt whether this is a peculiar failing of my own.

The whole matter is, of course, at bottom one of formal logic, and the necessary foundations are fully set forth by Russell and Whitehead in those sections of *Principia Mathematica* which deal with the ancestral relation (*R: see Pt. II, Sect. E, *90–*97, in Vol. i; also Introd. sect. VII and Appx. B in the second edition). No doubt, most of what is significant in the present essay could be expressed in their symbolism by any one sufficiently trained to its use. This, however, I am not; nor do I know whether full symbolic treatment of my argument would result in any practical convenience. Perhaps it would not be possible to say till the experiment had been tried. Meanwhile, I am acutely conscious that, compared with what the method

might achieve in abler hands, the present attempt is as *barbara celarent* to the modern logic of Peano and Wittgenstein.

I wish at the outset to make it clear that there is nothing esoteric or mysterious about my so-called Calculus: it aims at nothing but defining and making precise for formal use the logical rules which textual critics have always applied. It is quite incapable of producing any results that could not have been attained by the traditional methods; only it aims at achieving them with less labour and greater certainty. Perhaps its chief merit—if it has any at all—will be found in the endeavour to give precision to terms and modes of inference which are frequently employed with quite astonishing looseness. The working of it out has done so much to clear my own mind on the subject, that I cannot but hope that its study may be of some assistance to others.

The Calculus was not constructed *in vacuo* out of mere superfluity of naughtiness, but grew out of an attempt to determine the relation of the manuscripts of the Chester Plays, and the present essay began as a section of an introduction to the pageant of Antichrist in that cycle. It soon, however, became disproportionate, and now appears in separate form. I hope before long to publish my edition of the play, in which the method here described will find specific application.

It may be well to add that I am aware that about the middle of the eighteenth century Lagrange and Euler evolved a branch of mathematics known as the Calculus of Variations. It does not touch the problem discussed in the following pages.

Miss St. Clare Byrne has very kindly read the proofs for me.

CONTENTS

THE
CALCULUS OF VARIANTS

General Notions: Descent and Variation

If we exclude the possibility of memorial transmission,[1] all manuscripts of a given work are derived (by transcription) from a single original.[2]

The whole collection formed by the original together with all its descendants, in the particular relation in which they stand to one another, constitutes a family,[3] which, like other families, has a genealogical tree. Such a tree is the sum of all the lines of descent of the various manuscripts; a *line of descent* being a series whose consecutive terms are

[1] 'Memorial' is a better and generally rather wider term than 'oral'.

[2] In order to simplify exposition so far as possible I have deliberately narrowed the field explicitly covered. To have included memorial transmission would have necessitated somewhat different and more complicated definitions. On the other hand there is no need to exclude dictation, which is a mere incident of transcription. It may affect the character of the variants, not the principle of variation. Of course, print can be substituted for manuscript, again without alteration of the principles involved, though in practice the problems that arise are generally different. There seems no need to exclude from 'transcription' revision of the work by the author or another, but it could easily be done by formally postulating that such a recension constituted a different work. The case has not been explicitly considered in what follows.

[3] The term 'family' is often applied in a merely extensional sense to mean either the manuscripts of a work generally, or those of a particular branch. Here, however, it will always be used to include the genetic relation. Of course, if the inferential manuscripts (in the sense later defined) are specified, the relation is given, since they are merely the formal expression of that relation.

linked together by the relation of parent and child (exemplar and transcript). Normally the lines of descent are divergent in a downward, convergent in an upward, direction.[1]

In practice, however, we seldom have immediate knowledge of a whole family. What we find given is a set of *extant manuscripts* (two or more of which may belong to one line of descent, but which are more often not directly linked by the ancestral relation) from whose resemblances and differences we are able, by a logical or quasi-logical process, to infer the former existence of a number of what may be called inferential manuscripts. An *inferential manuscript* is a node of the genealogical tree, a point at which some line of descent branches.[2] Of course, the farthest that this process of inference can take us is back to the archetype of all the extant manuscripts. This may not be identical with the original

[1] This distinguishes the genealogy of a manuscript (or any parthenogenic) family from that of a human family (or any in which sexual generation obtains). In the former the genetic relation is always one-one or one-many, in the latter many-one or many-many. We have, however, in the case of conflation, a phenomenon in manuscript genealogy analogous to sexual generation, and giving rise to a many-one relation. Conflation is outside the purview of the present essay, but a few remarks on the subject will be found in Note B.

[2] An inferential manuscript is the latest exclusive common ancestor (as subsequently defined) of some group of extant manuscripts. I prefer the term 'inferential' to the more familiar 'hypothetical' because this latter has often a wider extension than is here desirable. We are, namely, at times able to conjecture the existence of hypothetical manuscripts that are in fact internodal (or ultranodal) points, intermediate between (or anterior to) extant or inferential manuscripts, but which do not themselves mark divisions in the line of descent. I doubt, however, whether the inference in these cases is strictly logical, or, at least, whether it is based on evidence of which the calculus can take account. Be this as it may, I have deliberately excluded such manuscripts, often including the 'original', from the definition of inferential manuscripts, relegating them, however regretfully, to the limbo of what I have called the potential.

postulated at the start (in practice it probably seldom is): but not only can the methods here contemplated take us no farther, they cannot even throw light on the question whether anything lies beyond. Connected with this ascertainable class of extant and inferential manuscripts, there is, of course, an indefinite number of others which probably once existed but whose identity can now be but seldom, and then only vaguely, apprehended. These may be called *potential manuscripts*. They have no interest for us here beyond the fact that the discovery of a new extant manuscript will generally raise certain of them to inferential rank.[1] We shall, therefore, define the *family* as consisting of the set of all extant manuscripts together with their archetype and the other inferential manuscripts needed to explain and express their mutual relation. Should it ever be desirable to make more explicit the distinction between the family as here defined and the wider conception with which we started, the former may conveniently be styled the *logical*, the latter the *potential family*.

In connexion with the genealogy of manuscripts several notions require definition.

By *ancestor* of a manuscript we mean any earlier manuscript in the same line of descent. It should be observed that 'ancestor' by itself is indefinite; we cannot in general speak of *the* ancestor, but only of *an* ancestor, of a manuscript. The notion becomes definite, however, when we speak of

The *latest ancestor* of a manuscript, which is, of course, its immediate parent.

Similar notions apply to groups of manuscripts, but the indefinite form is so unimportant that it is best disregarded, and we define the *common ancestor* of a group as the latest manuscript which is an ancestor of every member of the group, that is the

[1] Their possible existence will always be ignored in formal discussion.

latest manuscript common to the several lines of descent. It should be observed that any and every group of manuscripts selected from a family has of necessity a common ancestor, otherwise (by our original postulate excluding memorial transmission) its members could not all preserve the same work.

The most important notion of all is that of the *exclusive common ancestor* of a group, that is, the latest ancestor that is common to the group and to no other extant[1] manuscript. This, it will be observed, is not something different from, but a particular case of, the common ancestor. It follows that it does not always exist for any particular group; but at the same time the common ancestor can always be made the exclusive common ancestor by adding to the group the other manuscripts derived from it, where these are known.

Members of one family but of different lines of descent are called *collaterals*. Any group of manuscripts of a given work will therefore be of one of three types. It will be an *ancestral group* if the manuscripts it comprises belong to a single line of descent, that is, are all linked by the ancestral relation. It will be a *collateral group* if the manuscripts all belong to different lines of descent. Lastly, it will be a *mixed group* if it is neither purely ancestral nor purely collateral. A collateral group may, of course, include or consist of inferential manuscripts, and the extant manuscripts of a work may form a mixed group. If the collection of all extant manuscripts is

[1] The qualification is formally necessary, since otherwise we could not in general speak of the exclusive common ancestor of a group of extant manuscripts alone, which is generally just what we want to do. At the same time it is not intended to confine, and does not confine, the group to extant manuscripts. It is often convenient and quite legitimate to speak of the exclusive common ancestor of a group of, or including, inferential manuscripts; for in such a case these are really no more than symbols for the groups of extant manuscripts derived from them.

collateral, it may be called a *terminal group*, that is one consisting of manuscripts each of which is the end of some line of descent. All *terminal manuscripts* are both extant and mutually collateral, but neither extant nor collateral manuscripts are necessarily terminal.

Given a number of manuscripts of a work, which we will call A, B, C, D, . . ., their common ancestor and their exclusive common ancestor may for convenience be written A'ABCD... and xA'ABCD... respectively. We may also, if we so desire, use the symbols A and xA by themselves to mean respectively the common ancestor and the exclusive common ancestor of some group in question.[1]

Again, given xA'BC, say β, and also xA'ABC (i. e. xA'Aβ), say α, we may express these data in the single formula xA'A(BC). Or, given xA'CD, say γ, and also xA'ABCD (i. e. xA'ABγ), say α, from which A, B, and γ are independently derived, we may write xA'(A)(B)(CD). On the other hand, if, in the latter case, we had xA'AB, say β, we should, of course, write xA'(AB)(CD).[2]

This simple convention of putting

$$xA\text{'ABC} + xA\text{'BC} = xA\text{'A(BC)}$$

enables us to express the relation of any number of

[1] The word 'common' only serves to indicate that we are speaking with reference to a group and not an individual; when therefore the group is explicit it becomes superfluous, and is consequently dropped in the symbolism. In using the symbols by themselves, however, it should be remembered that they are only strictly applicable to groups.

[2] For the definition of independent derivation see below, p. 7. It would occasionally be convenient to write xA'AB(CD), where 'AB' should mean '(AB) *or* (A)(B)', but it is doubtful whether the occasional convenience of an indeterminate formula would compensate for the confusion its introduction might cause. I shall throughout use roman capitals to indicate extant manuscripts and small Greek letters to indicate inferential ones. In the few cases where it is necessary to distinguish between manuscripts and their readings, I shall indicate the latter by italic capitals.

manuscripts in symbolic form. Let us suppose, taking the example I shall use throughout, that a work is preserved in six extant manuscripts, namely A, B, C, D, E, and F, and that no two of these belong to the same line of descent.[1] Then, if, for example, there exist only xA'EF, xA'CD, and xA'CDEF, besides (of necessity) A'ABCDEF, we can completely define the family by the formula $(x)A'[A][B][(CD)(EF)]$. Here we write '(x)' instead of 'x' to indicate that it is only significant for the several sub-groups, for xA has no meaning in connexion with the sum of extant manuscripts. The slight formal distinction serves to indicate that we are considering a comprehensive relation: every formula beginning with '$(x)A$' defines a complete family, one, that is, comprising all extant (and consequently also all inferential) manuscripts.[2]

We may occasionally wish to assert the existence of xA of some group in respect to some larger group which, however, does not include all extant manuscripts, without implying anything as to those excluded. This may be done by writing, for example, $(CDEF)xA'EF$, which confines the field of the statement to the group CDEF among extant manuscripts, and leaves open the question whether $A'EF$ is also an ancestor of A or B or not.

It remains to observe that derivation is of two types, independent and successive. In one sense, and in connexion with particular groups, this is, of course, obvious. Derivation in the line of descent is necessarily successive, while any number of manu-

[1] To this condition of collaterality I shall return later; see p. 22 and Note A.

[2] For the sake of clearness, and for convenience of reference, a number of typical families of six manuscripts are exhibited diagrammatically on pp. 60–1, each accompanied by the formula that defines it. The families represented are, of course, only a selection from those theoretically possible. For brevity I shall speak of the formula as being, not merely as defining, the family.

scripts are independently derived from their immediate parent. But there is a less obvious, and derivative, though for our purpose more important, sense, in which the terms may be applied to whole families or even to collateral (especially terminal) groups. And here it should be observed that even the independent derivation of several manuscripts from their immediate source is successive in so far as a child succeeds its parent, while without some independent derivation no collaterals could come into existence. It follows that the definitions will depend on degree. *Independent derivation* is found throughout any collateral group for no selection from which does xA exist; that is, in the case of our six terminal manuscripts, only in the family $(x)A'(A)(B)(C)(D)(E)(F)$, in which succession is reduced to a single generation (i. e. genetic step). *Successive derivation* is the antithesis of independent but is less easily defined. It might appear sufficient to recognize as successively derived any family in which there were never more than two manuscripts independently derived from a common source. This, however, would not give a unique result, such as is desirable. We can obtain this by adding the condition that, of each pair of independently derived manuscripts, one at least shall be terminal. This is satisfied only by the family $(x)A'A\{B[C(D\overline{EF})]\}$.[1] For this, however, an equivalent and preferable definition is to be found in the fact that all the inferential manuscripts form an ancestral group. It is to this type, therefore, that we shall confine the term successive derivation. The looser type resulting from the definition first considered, and satisfied by $(x)A'\{AB\}\{C[D(EF)]\}$ and $(x)A'[(AB)C][D(EF)]$ and various other families, may be described as *quasi-successive.*

[1] It is, of course, also satisfied by $(x)A'\{[(\overline{ABC})D]E\}F$, but the two families are identical so long as A, B, C, ... remain variables.

Lastly in those cases which, without being purely independent, involve the derivation of at least three manuscripts, extant or inferential, from a common parent, such as the families $(x)A'(AB)(CD)(EF)$ and $(x)A'A\{B[(C)(D)(EF)]\}$, we may recognize the derivation as *quasi-independent*. The importance of the notions of independent and successive derivation lies in their relation to the corresponding forms of variation and divergence.[1]

The process of transcription is characterized by variation, and it is only in the process of transcription that variant readings arise.[2]

Such variation may be assumed to be universal, every transcription introducing some variants. This is obviously not necessarily true, but it agrees with experience in all but the shortest texts. Moreover, an operation that produces no effect may safely be ignored, and, should there be such a thing as an absolutely faithful transcript, we shall be led into no error if we treat it as identical with its exemplar. Most variants are spontaneous, that is to say that they are not in any way conditioned by variation in

[1] A few formal antitheses, out of many, may be noted. If derivation is purely independent, then, in the formula defining the family, there are the maximum number of brackets, these are all of the same order, and there is only a single generation; if derivation is purely successive, then there are the maximum number of brackets of different orders, no two pairs are of the same order, and there are the maximum number of generations, namely one less than the number of terminal manuscripts.

[2] This is not historically true, but it is a convenient and innocent assumption. Many variants in extant manuscripts have arisen through an alteration being made in an ancestor after the original scribe had completed his work. In such a case, transcripts made before the alteration will have one reading, those made after it another. But in order to render the statement in the text rigorous we only need to postulate that the alteration of a manuscript is equivalent to transcription, and, therefore, that the manuscript in its original state is not identical with, but the parent of, the same when altered.

the exemplar: on the other hand some are so conditioned, since a slip in one transcription often leads to emendation (correct or not) in the next. But we may safely assume that in no transcript are all variants thus predetermined; indeed, this almost of necessity follows from our former assumption. There also follows from it, at least in suitable cases, another and more extreme inference, namely, that, of the variants introduced in any transcript, some will persist through subsequent transcriptions, while others will undergo further variation. Since, in any transcription, only a small proportion of the readings undergo variation, the former part of this proposition will be readily allowed. The latter part is less obvious, but it will be observed that the variations introduced in the course of any transcription themselves form a textual field, over which, if it is sufficiently extensive, the assumption of universal variation will be operative. Moreover, the principle of predetermination will make this field more particularly subject to variation.

We require then, the following postulates:

Universal variation, namely, that every act of transcription introduces some variants;

That *spontaneous variation* is more widely effective than *determined variation*, and consequently that the variants introduced in any transcription are never all predetermined;

Persistence of variation and *variation of variation*, from which it follows that, of the features peculiar to any manuscript, provided they are sufficiently numerous, some are transmitted unaltered to its descendants while others are further modified.[1]

It should be observed that the term 'variation'

[1] Critics have sometimes tacitly assumed the further postulate of *constant variation*, namely, that every transcription introduces approximately the same number of variations in any given text. This is quite contrary to experience and leads to erroneous results (see Note C).

is used, strictly speaking, in two somewhat different senses, or at least is applied to two different cases. There is the variation of a descendant from an ancestor, and there is the variation of two collaterals from one another. The former may be called *vertical variation*, the latter *horizontal variation*. The former is fundamental, the latter derivative; for, of course, the variation between two collateral manuscripts is merely the effect (observable if they are extant) of the variation of one or both of them from their source. Horizontal variation is the datum, vertical the end, of textual criticism. It may be noted that horizontal variation always implies vertical variation in at least one line of descent; but vertical variation only leads of necessity to horizontal variation if it occurs within the limits of the logical family. In the complete potential family all lines of descent may pass through the manuscript in which variation arose. In other words, the readings of any collateral group are evidence of the reading of the archetype only, not of any earlier manuscript: which is obvious—though it seems to be sometimes forgotten.

Just as, in any family tree, different lines of descent are seen to be divergent in a downward direction, so the text, in any line of descent, becomes increasingly divergent both from the original and from that of any other line of descent, measuring divergence by the number of variants. This is presumably always true. We might proceed to argue that the number of variants between a manuscript and any ancestor was the sum of all the variants introduced in the intervening transcriptions, and that the number of variants between any collaterals was the sum of the variants introduced in the transcriptions intervening between them and their latest common ancestor. But this would only be true so long as the variants introduced were themselves divergent. This is

not always so. A variation in transcription may accidentally, and often does intentionally, restore the reading of an earlier ancestor. Also two independent transcriptions may alter a particular reading in the same way. In either case, the second variation, instead of increasing the divergence of the texts, reduces it. Thus, by the side of the normal *divergent variation*, we must recognize, in successful emendation and in the chance coincidence of error, two forms of what may be called *convergent variation*.[1]

Horizontal variation gives rise among collaterals to *grouping*, that is, to the arrangement of the manuscripts into groups according as they agree or differ in respect to particular readings. By a grouping, we understand a list of all the extant manuscripts (or of some selection of them) divided on this principle into two or more groups each of one or more manuscripts. But since it is only by a stretch of language that a single manuscript can be called a group, we may describe as *true groupings* those including at least two groups each of two or more manuscripts, and we shall find in the sequel that these alone are significant. We shall also see later on that the only groupings that can be regarded as *fundamental* are those that divide the manuscripts into two groups only. In such groupings we may speak of the two groups as the two sides of the grouping, and each as the complement of the other. More generally, the complement of any group is, of course, the group of all the other manuscripts in question.

Different variants will group the manuscripts in different ways. Should it be found that all possible arrangements occur with much the same frequency the grouping in general may be described

[1] With divergence considered, not in relation to the text as a whole, but as the degree of variation in particular variants, we shall be concerned later (see p. 30).

as *random*, and, of course, no inference as to the relation of the manuscripts is possible. To be of any evidential value the groupings must be *constant*, certain arrangements occurring to the exclusion of others with at least such regularity as to suggest that exceptions may be due to chance. When the groupings are constant, some inferences can always be made, but the results will be contradictory unless the groupings are not only constant but also *consistent*. The conditions required for consistency are not altogether easy to define formally, but for fundamental groupings they appear to be satisfied if, and only if, given any two constant groups, either these or their complements are either mutually exclusive or one wholly includes the other. (Groups are, of course, constant or consistent if they occur in constant or consistent groupings.) The rule comes to this, that while one or more manuscripts may pass from one side of a grouping to the other without rendering it inconsistent, those on opposite sides must not exchange places.

The grouping of the manuscripts may be considered either with reference to a particular variant, or generally with reference to several, or all, variants. The generalized grouping is, of course, the sum of the particular groupings, but it is of a much more complicated nature, since, not only are the constituent (fundamental) groups no longer confined to two, but their relation is no longer one of simple opposition. It is clear that we shall require a symbolism for variation in some ways parallel to that already adopted to express ancestry. But the development of this requires to be dealt with in greater detail and must be postponed to later sections of this essay.[1]

Meanwhile the parallelism just mentioned between variation and descent suggests a very important observation. It is, namely, necessary to distinguish

[1] See in particular p. 23, and further p. 44.

clearly between two different meanings of the term 'group'. The groups we have just been considering are what may be called *variational groups*, that is merely groups of manuscripts having certain readings in common. Since it is such groups that will mainly occupy our attention, I shall by *group* always mean a variational group unless some other is expressly indicated. But besides these there are *genetic groups*, or branches of the family tree, characterized by the possession of an exclusive common ancestor.[1] The two are, of course, related. Thus, if the manuscripts A, B, and C constitute a genetic group, this will give rise to variants in which ABC will be opposed to DEF. On the other hand, if we find that the variants habitually divide the manuscripts into the two groups ABC and DEF, then these will be significant constant groups; but, though both may also be genetic groups, that either ABC or DEF should be such will suffice to account for the facts. The process of determining the relationship of the manuscripts consists in inferring from the variational the corresponding genetic groups.

It is the object of the Calculus of Variants to facilitate this process by substituting, so far as may be convenient, the use of symbols and formal rules for the continuous application of reason, thereby not only economizing mental effort, but avoiding, it is hoped, certain confusions of thought which, as experience shows, are liable to occur.

[1] At first sight it might appear sufficient to postulate the existence of A'ABC as the condition of ABC forming a genetic group. Certainly, such a group would be, in some sense, genetic. But it would not be a complete genetic group, and it is necessary to include completeness in the notion, since otherwise any selection from the manuscripts would form a genetic group. The definition adopted for xA renders it unnecessary to include inferential manuscripts in order to secure completeness.

Recording Variants

Given once more our six extant manuscripts A, B, C, D, E, and F, it will always be possible to quote their variant readings according to some such formula as

xyyx ABC : *yxxy* DEF, or
xxyy AB : *xyyx* CD : *xyxy* EF,

in each case giving the words that replace one another in the different groups.[1] The most frequent formula will very likely be of the type

xyyx ABCDE : *yxxy* F,

which suggests the need of some symbol to indicate 'the rest'. Putting Σ for 'the sum of the unspecified manuscripts', the last-mentioned formula becomes

xyyx Σ : *yxxy* F.

It will be best, as a rule, to reserve Σ for the largest group in any formula;[2] thus we shall continue to write the first pair above as before, but we shall have, for instance,

xyyx Σ : *yxxy* EF, and
xxyy Σ : *xyyx* DE : *xyxy* F.

The meaning of Σ is defined in relation to the manuscripts that are generally available throughout the text. Should one or more of these show lacunas, then, in the passages affected, some modification of the symbol must be used. For instance, if a stanza, say, is omitted in E and F, then, in quoting variants between the other manuscripts in these lines, we must replace Σ by the qualified symbol Σ – EF, more conveniently written Σ_{EF}. Similarly, should there be

[1] Where minor differences of spelling and so forth are neglected, the words (unless normalized) should be quoted in the exact form in which they occur in the first of the manuscripts cited.

[2] Occasionally, however, in discussion it is convenient to designate some smaller group by Σ. It does not generally seem worth while to substitute Σ in the case of two manuscripts only, even when these constitute the largest group.

a fragmentary manuscript Z available in parts, then, where this is so, Σ must become Σ + Z, or Σ^{Z}. In connexion with this latter case a small formal point deserves mention. Should the reading of Z differ from that of all the other manuscripts, it might, though correct, be slightly misleading to write

$$xyyx\ \Sigma^{Z} : yxxy\ Z,$$

and no confusion can arise if we put instead

$$xyyx\ \Sigma : yxxy\ Z.$$

Whenever, therefore, Z is explicitly mentioned Σ may be substituted for Σ^{Z}, and we shall, for instance, write

$$xxyy\ \Sigma : xyyx\ EZ : xyxy\ F.$$

Should occasion arise, such symbols as Σ^{Z}_{EF}, or Σ^{YZ}_{C} may, of course, be used to indicate the unspecified manuscripts of the collections ABCDZ and ABDEFYZ respectively.

What has been said so far applies where variants are quoted without reference to any particular text, and must, of course, be followed when handling formulas for the purposes of discussion. A few words may be added on the conventions best suited to recording variants in connexion with a printed text. A common practice is always to give first the reading of that text, and to separate it from what follows by a single bracket. Thus, for example, we might find

$$xyyx] \Sigma : yxxy\ EF, \text{ or}$$
$$xyyx] A : yxxy\ \Sigma.^{1}$$

In the former of these examples we may omit the Σ, since the reading of the text may be assumed to be that of all the unspecified manuscripts, and write simply

$$xyyx] yxxy\ EF$$

without ambiguity. On the other hand we cannot

[1] In these formulas it would be more usual to write 'so Σ' and 'so A', but there is no necessity to do so.

similarly omit the A in the second example, since where there is no specified group Σ can have no meaning. And even were we to replace Σ by BCDEF the omission of A would still be undesirable.

This applies to cases in which the printed text is a critical one. Where it reproduces exactly[1] a single manuscript, the fact of a reading appearing before the bracket is equivalent to the specification of that manuscript, and in that case Σ may be used without ambiguity even if the letter indicating the particular manuscript be omitted. It is, however, doubtful whether anything is gained by the omission.

The reading before the bracket, 'taken down' from the printed text, is called a 'lemma'. Lemmas must agree exactly with the text.[2] Some difference of practice exists as to 'taking down' punctuation along with the reading. In the case of a critical text, in which the punctuation is the editor's and therefore of no critical value, it is best neglected altogether. On the other hand, where the text reproduces exactly (or approximately) a particular manuscript, it may often be important to record differences of punctuation found elsewhere. In this case it is sometimes held essential that, since a variant may consist in the absence or addition of punctuation, it should be 'taken down' in all cases. This is a mistake. Punctuation need only be 'taken down' when it is itself in question. When that is so, the variant consists in the difference of punctua-

[1] Or generally, provided the exceptions are clearly marked.

[2] A slight exception to this generally rigid rule is that diacritics used in the text need not necessarily be retained in the lemma. For example, if certain letters are printed in italic to indicate the expansion of an abbreviation, it may be advisable to neglect the distinction in 'taking down'. So with brackets indicating mutilation and so forth, which tend to cause confusion in the lemma. An editor must settle these points for himself, but he must settle them consistently and with appreciation of what is involved.

tion recorded; the preceding word is merely added to indicate the position; no ambiguity can arise.

Some editors omit lemmas altogether, relying on the sense to show which word or words in the text the variant is intended to replace. In the hands of an editor of approved competence—*rara avis in terris*—the practice is unobjectionable, but it demands a degree of skill and vigilance to say the least uncommon, and is certainly not to be generally recommended.[1] Instances could be cited in which an *apparatus criticus* has been rendered largely worthless through the lax use of this method. In particular the warning is desirable that, where the variant consists of an omission or an addition, the context on both sides should be given. Further, if variants of punctuation are taken into account at all, it will here be necessary to regard a point as an integral part of the preceding word, and in all cases to quote the one along with the other, else there will be no logical means of recording its absence.[2]

It should be added that no directions can be of universal application. Editors will always have to adopt special conventions to meet particular needs.

A matter of some importance, that falls for discussion here, is the *degree of collation*, that is, the minuteness of the variants of which notice is taken. Needless to say, this should be constant throughout—so far as possible. Otherwise it is a matter of choice. If we confine our attention to the more important variants, we can be fairly certain, provided we are dealing with the work of a naïve scribe, that the readings are meant to be those of the exemplar, and are evidence of the descent of the manuscript in which they occur. On the other hand, if the work

[1] I may mention that I have attempted the method myself and abandoned it owing to the difficulty of avoiding ambiguity.

[2] Of course, some *ad hoc* device of annotation may be adopted, and this may be the best means of meeting the difficulty, if it does not arise too often.

is a short one, we risk limiting the field too much to eliminate the operation of chance: moreover, there are other dangers connected with conflation which will be touched on elsewhere.[1] If, however, we make our collation very detailed, we are met with difficulties of another sort. For, whereas, in major matters, a scribe will, as a rule, follow his exemplar, in the minor points of spelling and grammatical form he will be largely led by his own fancy. Consequently, the more minute we make our collation, the greater the number of non-evidential variants we shall be recording, and the greater the risk of chance coincidences between manuscripts. It is obvious that transcripts of different exemplars by the same scribe will show marked resemblance in the minor readings and marked differences in the major; whereas transcripts of the same exemplar by different scribes may agree almost throughout in important matters and yet will differ widely in detail. The degree of collation desirable must be decided in every case in relation to the character of the work contemplated.

Types of Variants

The formulas expressing the variation of the manuscripts in respect to particular readings conform to a number of definite *types*. There are, to begin with, two main classes, the simple and the complex. In *simple variants* the formula defines two alternative groups, to one or other of which every manuscript belongs. In *complex variants* the groups are more than two in number. The simple class comprises two types: type 1, in which one group consists of a single manuscript, and type 2, in which either group consists of two or more

[1] In Note B.

manuscripts. The complex class comprises types 3, 4, 5, ..., according as there are 3, 4, 5, ... groups in the formula. Thus the number of types is the same as the number of manuscripts. For example, in the case of our six terminal manuscripts we shall find the following types:

Simple	Type 1	Σ : A, Σ : F, &c.
	Type 2	Σ : AB, Σ : EF, ABC : DEF, &c.
Complex	Type 3	Σ : A : F, Σ : DE : F, AB : CD : EF, &c.
	Type 4	Σ : A : E : F, AB : C : D : EF, &c.
	Type 5	AB : C : D : E : F, &c.
	Type 6	A : B : C : D : E : F.

It will be observed that several different formulas are possible under each type except the highest: these may be called the different *forms* of the type. They comprise all the different arrangements into which the manuscripts can fall, and any form may, and usually will, be exemplified in a number of instances.

Since every manuscript contains variations from its immediate source, any reading supported by one manuscript alone may have originated in that manuscript, and such a reading therefore cannot, without further analysis, throw any light on the relation of the manuscripts of the collateral group. To establish such a reading as original it would have to be shown, not only that the reading was correct, but that it could not be due to emendation.[1] To prove either of these is strictly impossible, and though in individual cases the probability may be great, and

[1] Even if this were admitted, to prove for the other manuscripts a common and independent derivation, it would be necessary to assume that their common error had not arisen independently in the course of various transcriptions. This assumption we do habitually make, and although it is not necessarily correct in any individual case, without it no inference as to the relation of the manuscripts would be possible.

repeated cases give rise to moral certainty, this is hardly a mathematical notion, and therefore finds no place in the calculus.[1]

It follows, therefore, that only those variants which give rise to at least two groups of more than one manuscript each can be described as (genetically) *significant variants*. And only those which give rise to groups all of which are of more than one manuscript can be described as *completely significant*. By *significant groups* we shall understand true groups (i. e. of two or more manuscripts) arising from significant variants.

It will be observed that, if the number of manuscripts and types is n, types 1, n, and $n-1$ can never, as they stand, be significant, and that the highest type that can be completely significant is $n/2$, if n is even, or $(n-1)/2$, if n is odd. Also, that the number of possible forms of type 1 is n, and of type n is 1. The possibilities of the intermediate types may be left to mathematicians to determine.

It may be asked why the question of significance has been allowed to divide the class of simple variants into two types, whereas the distinction has not been carried through the complex class. The reason is that the point has not the same importance in the higher types as in the lower. For, though a type-1 variant can never be significant, we shall shortly see that where more than two groups arise the reading of a single manuscript may show an affinity that will enable the variant to be reduced to one completely significant.

[1] Though the matter lies beyond my present theme I may point out a curious difficulty that arises when we attempt to infer manuscript relation from supposed originality. To show that a reading is original two main lines of argument are available: that the reading is itself satisfactory, and that it explains the origin of the erroneous alternative. But, as a rule, the easier it is to explain how an error arose, the less valid the assumption that it only arose once. Thus the more likely it is that one alternative is correct, the less certain it is that the other points to common derivation.

Principles of Variation

From what has been said in the previous section it will be clear that, where three manuscripts only are concerned, no merely formal process can throw light on the relationship between them. Either the readings will be all divergent or else the variants will be of type 1, and since, in the latter case, the reading of the single divergent manuscript may always (theoretically at least) be unoriginal, it will never be possible to establish a common source for any pair of manuscripts to the exclusion of the third. Given three manuscripts, therefore, it is impossible either to prove or to disprove independent derivation. This fact, which I call the *ambiguity of three texts*, we shall find meet us at every turn of the discussion, and it largely determines the nature of the calculus.

But though type-1 variants, whatever may be their individual interest, are of no use for our present purpose, either the absence of higher types, or the absence of particular forms of the lowest, may be of evidential value. Suppose that, in the text preserved in our six manuscripts, none but type-1 variants occur. At first this might seem to imply that all the manuscripts were independently derived from a common ancestor, but, owing to a complication not unlike that mentioned above, the correct inference is that at least all but one of the manuscripts are so derived. A couple of examples will make this clear. Suppose there exists xA'AB, say α, then the writing of α will have introduced certain variants, and some of them will (according to the postulate of the persistence of variation) have survived in both derivatives, giving rise to type-2 variants of the form Σ : AB. Consequently the absence of such variants disproves the existence of xA'AB. Suppose, however, that there exists

xA'BCDEF, say β, then variants arising in β and surviving in all derivatives will give rise only to type-1 variants of the form $\Sigma : A$.[1] Consequently the absence of variants of higher type (or of significant variants generally) does not disprove the existence of xA of all the manuscripts except one.

It follows from the postulate of universal variation that in any collateral group all forms of type-1 variants must occur. Consequently the absence of one or more forms of this type is inconsistent with the assumption that our six extant manuscripts form a collateral group. This might seem to offer a means of eliminating ancestral elements from any collection and thus replacing our assumption by a logical demonstration. Unfortunately, a consideration of what has been said above shows that the converse is not necessarily true, that the presence of all forms of type-1 variants does not prove that the collection is collateral throughout. Thus the elimination of descendants cannot be effected by the sole use of the calculus, and the further consideration of the question must necessarily be postponed.[2]

Type-2 variants are a very different story. If we have a variant AB : CD, then one or other reading must differ from that of the archetype, and one or other group must be genetic : there can be no question of all four manuscripts being independently derived. Different forms of type-2 variants will divide up our collection in different ways, and these divisions will correspond to the ramifications of the family tree. It is clear that some symbolism can readily be devised that will represent the complete grouping afforded by the variants, and that the

[1] In this instance it should be observed that, while the appearance of $\Sigma : A$ proves that A is not identical with β, unless we know that our extant manuscripts are in fact collaterals (as we assume them to be), we cannot tell whether A is an ancestor or a collateral of β (cf. p. 51).

[2] See Note A.

correspondence between this and the ancestral symbolism already adopted will give the basis of such inference as can be drawn from variational in the direction of genetic grouping. Provided they are numerous enough, type-2 variants afford us all the evidence of which we can in the calculus make use. Not only so, but, although individual variants of higher type often look very tempting as a basis of inference, it is exceedingly difficult to devise any formal method of dealing with a number of them; the relationships are too complex to be readily amenable to rule.

When, however, we pass to the consideration of these variants of higher type a very important fact emerges. For, if we bear in mind that an actual variant, however complex, can only arise through variation in individual acts of transcription, and that at any point in the text a single act of transcription can only give rise to a single variation, it will be apparent that it is only such variation as we see in type 2 that is fundamentally significant.

It will be convenient to consider this matter more fully in connexion with the several varieties of variation, namely independent and successive, and for this purpose to develop somewhat our symbolism of variation. Thus, if, for instance, we constantly find the grouping $\Sigma : DEF$ and also the grouping $\Sigma : EF$, it is proposed to express this double fact by the compounded formula $\Sigma : D(EF)$. Similarly, if we find only the constant groupings $\Sigma : AB$ and $\Sigma : EF$ in variants of type 2 (that is, if such groups as CD and DEF are absent), we shall write $\Sigma : (C)(D)(EF)$. Thus, just as in the case of genetic groups we put

$$xA'DEF + xA'EF = xA'D(EF),$$

so now in the case of variational groups we put

$$\Sigma : DEF + \Sigma : EF = \Sigma : D(EF).$$

The parallelism of the formulas reflects the relation

between the two kinds of grouping: the former always implies the latter, though the latter does not necessarily imply the former. It should also be noted, as part of the parallelism, that since genetic groups mark single steps in derivation, only simple variants can be compounded.

Turning now to the varieties of variation, we observe that in any particular passage a single variation must, of course, arise at one particular point in the family tree. Further, when, in the course of the various transcriptions that generate the family, a particular passage is subjected to repeated variation at different points, these points may lie in the same or in different lines of descent. Thus a complex variant always implies that variation has taken place at more than one point,[1] and its nature will be determined by the relation in which those points stand to one another.

Variation being primarily vertical, successive and independent variation are notions that apply in the first instance to the relation between manuscripts and their sources. Thus, in the simplest cases, successive variation is seen where variation occurs in successive acts of transcription, that is, where a manuscript varies a reading which arose through variation in its parent; and independent variation is seen where variation occurs independently in two acts of transcription performed on one exemplar, that is, where two manuscripts both depart from the reading of their parent. More generally, we shall say that *successive variation* is multiple variation occurring

[1] Historically, multiple variation does not always lead to complex variants. If only two manuscripts are extant, there obviously can be no complex variants, yet both may have altered the reading of the source. Similarly, in a single line of descent, any number of variations may occur, and yet only the terminal reading survive. But since the calculus can only take cognizance of extant readings, the converse of the statement in the text is, for our purposes, also true, where enough manuscripts survive.

in a single line of descent, while *independent variation* is that occurring in different lines. Of course, we wish to speak of the variation between extant manuscripts, and there will be no objection to applying the terms as defined, in accordance with the manner in which the variants arose, so long as we are not tempted to confuse the derivation of the readings with the derivation of the manuscripts.[1]

Whether, and if so how far, it is possible, from the nature of the variants themselves, to ascertain whether they arose successively or independently, is a problem that will engage our attention when, in the next section, we come to discuss variational divergence. Meanwhile, assuming the nature of the variation to be given, let us consider the genesis of complex variants.

Taking first the case where variation is independent, let us suppose that there exist xA'AB, say α, xA'CD, say γ, and xA'EF, say ϵ, and that α, γ, and ϵ are independently derived from the archetype; assumptions expressed by the formula

$$(x)A'(\mathrm{AB})(\mathrm{CD})(\mathrm{EF}).$$

The only constant groupings in type 2 that can occur in such a family are $\Sigma : \mathrm{AB}$, $\Sigma : \mathrm{CD}$, and $\Sigma : \mathrm{EF}$ (i.e. $\Sigma : (\mathrm{CD})(\mathrm{EF})$ when compounded), due to variation in α, γ, and ϵ respectively, and it is merely through the chance concurrence of two of these variations that the complex grouping $\mathrm{AB} : \mathrm{CD} : \mathrm{EF}$ can arise. It is evident, therefore, that the correct way to regard this type-3 variant is as the product of two variants of type 2, though without knowing

[1] If we are given three *readings* $A : B : C$, and are told that B and C vary successively from A, what is meant, of course, is that C is derived from B, and B from A. But if the statement is made, in respect to a particular variant, about the *manuscripts* A : B : C, then what is meant is still the same, namely, that the reading of C is derived from that of B, and the reading of B from that of A, *not* that C is derived from B, or B from A.

which reading is original, there is no telling in which two inferential manuscripts variation occurred.

On the other hand, to take a case in which variation is successive, consider the family

$$(x)A\text{'}A\{B[C(D\overline{EF})]\}.$$

Here the forms of type-2 variants are different from what we found before, the compounded formula being $\Sigma : C[D(EF)]$. The variant $\Sigma : EF$ will result from variation in $xA\text{'}EF$, say ϵ, and $\Sigma : AB$ from variation in $xA\text{'}CDEF$, say γ, and the complex grouping $AB : CD : EF$ can only arise through chance concurrence of variation in γ and ϵ in succession.

Now consider the family $(x)A\text{'}[A][B][(CD)(EF)]$, in which we have only $xA\text{'}CD$, say γ, $xA\text{'}EF$, say ϵ, and $xA\text{'}CDEF$, say ϕ. The constant groupings will be $\Sigma : (CD)(EF)$, as in our first example. But here the complex variant $AB : CD : EF$ may be due to the concurrence of independent variation in γ and ϵ, or else to the concurrence of successive variation in ϕ and either γ or ϵ.

From this it will appear that independent and successive variation may be combined in a single complex. Thus, in the family

$$(x)A\text{'}[A][BC][D(EF)],$$

we might meet with the grouping $A : BC : D : EF$, due to independent variation in $xA\text{'}BC$ and $xA\text{'}DEF$, and variation in $xA\text{'}EF$ successive to that in $xA\text{'}DEF$. But it will be noticed that the same grouping would arise from variation in A, $xA\text{'}BC$, and D, when it would be independent throughout.

Thus while it is strictly true that every variant of complex type is the product of two or more simple variants, it is not possible (even given the family relation) to resolve variants into their factors without such a specific knowledge of their individual character as can only be obtained from a study of the actual

readings. It follows that in analysis the greatest importance attaches to the consideration of how the origin of a variant may reveal itself in the divergence of the readings to which it gives rise.

Before, however, passing (in the next section) to the further consideration of this problem, it will be necessary to discuss the treatment of *defective* and *redundant grouping*, since, though they are of no great consequence in themselves, the former acquires considerable importance in connexion with the formal aspect of resolution.

Let us suppose the scribe of some manuscript, say F, to have omitted a passage, and, at a certain point in that passage, the group ABC to have one reading and DE another. Then it is pretty clear that the absence of any corresponding reading in F is of the nature of a variant, and that the formula will have to be Σ : DE : F, since we do not know with which group the exemplar of F agreed, nor that F would not have varied individually had it reproduced the passage. It is, perhaps, less obvious that, if the passage has been lost from F through subsequent mutilation, the absence of any corresponding reading is equally of the nature of a variant. Yet it will be observed that here our ignorance is the same, and that the cause of our ignorance is indifferent. In such cases we agreed to substitute Σ_F in place of Σ, and we shall consequently write the formula Σ_F : DE. The question will occur: What is the importance of distinguishing between Σ and Σ_F? Now, if, instead of Σ_F : DE, we were simply to write Σ : DE, we should imply a grouping ABCF : DE. But there may be a constant grouping ABCD : EF, and the two are, according to our definition, inconsistent. But we do not know that the reading of F's exemplar agreed with ABC rather than with DE, and we, therefore, have no warrant for assuming that the real grouping contradicted the grouping

ABCD : EF (with which ABC : DEF is perfectly consistent). Thus the importance of the distinction between Σ and Σ_F is that without it we risk the invention of anomalous groupings. At the same time, so long as we retain Σ_F, the variants in which it occurs are all in effect of complex type, and of a variety in which no resolution is possible. It is therefore highly desirable to restore Σ wherever it is legitimate to do so. This is always the case when the resultant variant is of type 1, that is, when (apart from the defection) the only divergent reading is that of a single manuscript. $\Sigma_F : E$ is a defective grouping which conceals one of the following, viz.

$$\Sigma : E : F, \quad \Sigma : EF, \quad \Sigma : E.$$

Of these alternatives we may always safely choose the last, since from it no inference can be drawn,[1] and consequently in place of $\Sigma_F : E$ we are always at liberty to write $\Sigma : E$.[2] On the other hand, whenever (apart from the defection) the grouping is significant, that is, whenever there are at least two readings each supported by at least two manuscripts, to substitute Σ for the qualified symbol must always alter the meaning in an effective manner. Whether there is any objection to this will depend on the resultant grouping. Given a constant and con-

[1] The objection to $\Sigma : E : F$ is that it assumes a specific variation of F at this point. There is no great harm in this; but, since every transcript is more often correct than not, the assumption is contrary to probability. The bearing of this will appear when we consider the evidential value of individual manuscripts. $\Sigma : EF$, a type-2 variant, of course involves significant grouping.

[2] Of course, if we knew that the reading preserved by E was original, it would be by no means indifferent whether we wrote $\Sigma : E$ or $\Sigma_F : E$, since in that case $\Sigma : E$ would imply the existence of xA'ABCDF, which might be quite untrue so far as the inclusion of F is concerned. Thus when we come to construct a critical text we must be careful not to reject the reading of E in the variant $\Sigma_F : E$ on the ground that in the variant $\Sigma : E$ it cannot be original.

sistent grouping ABC : DEF, it would be illegitimate to substitute an unqualified Σ in the formula Σ_D : CEF, since it would indicate the inconsistent grouping ABD : CEF; but, the creation of such groupings being the only danger of substitution, there would be no objection, in the same circumstances, to writing an unqualified Σ in the formula Σ_C : DEF. And it will be observed that we can always (on sufficient data) get rid of the qualification by writing the formula the other way round. Thus, in the case supposed above, Σ_D : CEF is equivalent to Σ_D : AB, and this may safely be written Σ : AB, since this is consistent with ABC : DEF. All which, of course, comes to the same as saying that where a manuscript is defective we are at liberty to assume that its readings, if preserved, would conform to the constant groupings found elsewhere.[1]

Redundant grouping can be dealt with much more briefly. We previously agreed that, in portions of the text for which an additional manuscript Z became available, we would substitute for Σ the qualified symbol Σ^z. In regard to this only two observations appear necessary. So far as the relationship of the original manuscripts is concerned we can simply leave Z out of consideration, just as we are forced to leave lost manuscripts throughout.[2] As regards the relationship of Z to the original manuscripts, evidence can, of course, only be obtained from those portions of the text where Z is present. In these, therefore, Σ^z simply becomes a new Σ, subject to precisely the

[1] Cases may arise in which, whatever reading we assume for the defective manuscript, the grouping will remain anomalous. But in these the anomaly is independent of the defection, and we must merely be careful not to assume a reading that will increase it.

[2] This does not mean that Z can throw no light on the relationship of the original manuscripts. Its position with respect to them may be such that its presence will resolve the ambiguity of the three (original) texts.

same rules as the old, and such a further qualification as Σ^z_F may be treated in exactly the same way as Σ_γ.

Resolution of Variants

We have seen that variants of complex types arise as the product of those of simple types, into which, therefore, they may, on sufficient evidence, be resolved. But in the examples we took to illustrate this fact we assumed that the relation of the manuscripts was known, and then considered how complex variants might arise. In practice, of course, we approach the problem from the other end, and seek to infer the relation of the manuscripts from the form of the variants. If, therefore, we are to resolve variants of complex types, it must be on the basis of the intrinsic nature of the readings, and not on that of any assumption regarding the derivation of the manuscripts. The question consequently arises, whether in fact complex variants do exhibit peculiarities of such a kind as to enable us to form any opinion as to their mode of origin. I do not think it will be questioned that to some extent this is so, but the discussion will necessitate judgements respecting the character of individual variants of a kind we have not hitherto contemplated. So far we have merely been concerned with variation, in which readings are treated as being simply either the same or different. It is now necessary to introduce the notion of *divergence*, which takes account not only of the fact, but also of the degree, of difference. At the same time, it is important to observe that the judgements in question do not involve the correctness or originality of the various readings, but only their relative resemblance.

We have previously met with the idea of divergence of lines of descent and the progressive

divergence of the text in different lines of descent. These may for convenience be called *derivational* and *textual divergence* respectively. Textual divergence we saw to consist in the generally increasing number of variants. It finds expression in the compounded variational formulas, and since, as we also saw, these take account of simple variants only, it is clear that textual divergence would be unaffected by the total absence of complex variants. The notion with which we are here concerned, namely *variational divergence*, which will always be meant when the term 'divergence' is used without qualification, applies only to complex variants, and to these only individually. It does not lead directly or formally to textual divergence, since this is the generalized notion of simple variation (as shown in the compounded formulas), but one might regard variational divergence as the natural compounding of simple variants in the manuscripts themselves. The only reason for introducing the notion specifically is in order to allow of an analysis that will make possible the resolution of complex variants.

In the case of variation we regarded it as indifferent whether A was said to vary from B, or B from A, and we wrote Σ (generally) for the largest group, without suggestion as to originality. So, in variational divergence, when a particular reading is supported by more manuscripts than any of its rivals (when, that is, it appears in collation as the reading of Σ) we are at liberty to regard it as the basis, and the others as divergent from it.[1] But it is, of course, evident that, behind this variational divergence, there lies what may be called *genetic divergence*, that

[1] In discussion, just as it is sometimes convenient to use Σ for a smaller group, it is permissible to speak of the divergence of a larger group. It is, of course, the order not the size of the groups that is important, and it is often convenient to use Σ, independently of size, to indicate the point of view.

is, the divergence of readings, not merely in the order of their likeness, but in the sense in which they have actually arisen. We shall not be explicitly concerned with this notion, but it is sometimes helpful to bear the distinction in mind.

In what cases do complex variants show such peculiarities as to enable us to form some opinion as to how they arose? or rather: In what cases do the readings possess a clear order of similarity? It is, properly speaking, the second of these questions that we must ask, since to say that one reading arose from another would involve a judgement of originality. And we must observe that it is not mere violence of variation that is in question. If one reading makes an army march four parasangs, another five, and another six, there is nothing in this to establish any order of similarity; the variations may have occurred in any manner. I think that it will be found that in all instances in which a complete order emerges the readings themselves (as distinct from the variant) are in a manner complex. They comprise, namely, two or more parts or elements, and it is these that are separately varied in the different manuscripts, rather than the reading as a whole. This will become clearer in the sequel.

The investigation of the types of divergence is so closely bound up with the possibilities of resolution that the two must be considered together. In what follows, it will, I think, help to make the argument clear, if I take concrete examples in illustration, but for safety's sake I will add, in parentheses after each example, a purely conventional symbolic formula, which shall express the same variational relation, but which, being otherwise meaningless, will run no risk of introducing irrelevant ideas.

We may begin our discussion of divergence and the possibilities of resolution by observing that certain complex variants are only *apparent*, being

due to mere chance convenience in the manner of record. For instance, it may be desirable to give a collation in the form

thou . . . art AB : *thou . . . is* CD : *he . . . is* EF
(*xy . . . yz* AB : *xy . . . zy* CD : *yx . . . zy* EF).

But here it is evident that the readings could be recorded with equal correctness in the form

thou Σ : *he* EF and *is* Σ : *art* AB
(*xy* Σ : *yx* EF and *zy* Σ : *yz* AB),

and in this case, therefore, the resolution of the original type-3 formula AB : CD : EF as

(Σ : EF) . (Σ : AB),

where each factor is of type 2, is obvious. It does not follow that this resolution preserves the full significance of the variants, but it preserves all that the calculus can deal with, and a study of the symbolic formulas suggests that any further inference we might be inclined to draw from the particular example might well prove incorrect.[1] It is only after apparent cases of this nature have been eliminated that the real problem of resolution presents itself.

Before passing on we may observe that this example also illustrates another point. If we start with the reading *thou . . . art*, then *thou . . . is* must be a variant of this, and *he . . . is* a further variant of *thou . . . is*, and the divergence will be successive. If on the other hand we start with the reading *thou . . . is*, then both *thou . . . art* and *he . . . is* must be direct variants of this and neither of the other, and the divergence will be independent. Consequently, if the grouping had been, for instance,

thou . . . is Σ : *thou . . . art* E : *he . . . is* F

[1] For example, we might be tempted to say that because the reading of CD is erroneous it cannot be original. But this is not so: it may very likely preserve an error of the archetype for which the more plausible readings of AB and EF are alternative emendations.

the divergence would have been classed as independent; but if it had been, for instance, either

thou . . . art Σ : *thou . . . is* E : *he . . . is* F, or
he . . . is Σ : *thou . . . is* E : *thou . . . art* F

the divergence would have been classed as successive. Thus the distinction between independent and successive divergence is seen to depend upon the point of view, and remains merely formal until the introduction of the notion of originality gives one point of view preference over another. Whence it also follows that if resolution can be effected in one case it can also be effected in the other.

Although, of course, the majority of complex variants are not apparent in the sense of the example we have been discussing, and are incapable of being quoted except in their own proper types, a consideration of that example may suggest a method of dealing with them. Take, for instance, the readings

composed AB : *composing* CD : *reposing* EF
(*xyyx* AB : *xyxy* CD : *yxxy* EF).

These can be quoted in no other form, yet the case is essentially similar to the former. But, though we cannot here immediately resolve the formula

AB : CD : EF as (Σ : EF) . (Σ : AB),

we can, by first considering only those readings that agree in the ending and differ in the beginning of the word, write the first factor as Σ_{AB} : EF, and then, by considering only those that agree in the beginning and differ in the ending, write the second factor as Σ_{EF} : AB. And this method is generally applicable even where there is no possibility of splitting up the readings into component elements. Thus, if we take the readings

say AB : *tell* CD : *sing* EF
(*xx* AB : *xy* CD : *xz* EF),

then, by first neglecting AB we can write Σ_{AB} : EF,

and next by neglecting EF we can write Σ_{EF} : AB, just as before. Therefore, a complex variant such as AB : CD : EF can always be represented as $(\Sigma_{AB} : EF) . (\Sigma_{EF} : AB)$, or as the product of some similar pair of factors, and it is upon the handling of these factors that possibility of resolution depends. This use of the qualified symbols Σ_{AB}, Σ_{EF}, &c. is, of course, slightly different from that with which we are already familiar. Here, in writing Σ_{AB}, we take the group AB as hypothetically absent in order for the moment to disregard it. The conditions under which Σ_{AB}, Σ_{EF}, and so forth, can be replaced by the unqualified symbol, are generally the same as before, but we shall find as we proceed that the fact that the manuscripts indicated are not really but only hypothetically absent allows us to take certain liberties which would not be generally permissible.

Here then we have what may be called the *principle of resolution.* It consists simply of taking the differences that go to build up the complex variants one at a time. The possibility or legitimacy of so doing follows, of course, from the fact that, as we saw before, all textual divergence is generated step by step through variation in individual manuscripts. But, since we are now tracing the process backwards, certain ambiguities arise which were not previously apparent, ambiguities that find formal expression in the presence of the qualified symbols. Whether the resolutions are effective or merely formal depends upon the possibility of removing these qualifications.

There is one class of cases in which this is always possible. From what has been said above it is clear that such a variant as $\Sigma : E : F$ becomes when factorized $(\Sigma_F : E) . (\Sigma_E : F)$. But in our previous discussion of defective groups we saw that in variants of type I the qualification of the symbol may always be omitted. It follows that $\Sigma : E : F$ may always be written $(\Sigma : E) . (\Sigma : F)$. Thus variants in which

there is only one true group can always be resolved at sight without further inquiry.[1] In other cases the qualification will not automatically disappear, and it is only where, through the operation of special conditions to be considered later, the qualification can be removed from all factors that an effective resolution can be obtained.

It will have already become apparent that we shall have to consider three main types of divergence, namely independent, successive, and indeterminate. But since the distinction between the first two of these depends merely upon the point of view, a more fundamental difference is that of *indeterminate divergence* on the one hand, in which no order of similarity can be established, and *determinate divergence*, in which the order is complete, on the other. Independent and successive divergence are then seen as varieties of the latter. Of these we will now take examples and consider what can be achieved in the way of resolution. We shall, of course, have to take completely significant complex variants, and these, for our collection of six manuscripts only, will be of the form AB : CD : EF. In this I shall use the symbol Σ to indicate the point of view from which I wish to regard it, writing Σ : CD : EF or Σ : AB : EF or Σ : AB : CD at will.

The formal nature of the distinction between successive and independent divergence makes it convenient to take a single example to illustrate both determinate varieties. Let this be afforded by the readings

AB *To you I tell* (*xyyx*)
CD *To you I say* (*xyxy*)
EF *I say to you* (*yxxy*).

Here the order is completely determinate, the first

[1] But see p. 40, note 2.

and last readings resemble one another less than either does the second, that is, the second is intermediate between the extremes of the other two.

Let us now look at this from the point of view of CD, writing the formula $\Sigma : AB : EF$. If the reading of this group is original, those of AB and EF must be independently derived from it; the reading of AB cannot possibly have arisen through that of EF, nor the reading of EF through that of AB. Thus the instance is one of *independent divergence.* And, since the two variations arose separately, each might have occurred without the other, and it is, therefore, evident that the complex variant $\Sigma : AB : EF$ is merely the product of the two simple variants $\Sigma : AB$ and $\Sigma : EF$, into which it may consequently be resolved. A formal proof of this, which makes no assumption as to originality, is as follows. $\Sigma : AB : EF$ when factorized becomes $(\Sigma_{EF} : AB) . (\Sigma_{AB} : EF)$. But, in the first factor, EF is not really absent, and since its reading is closer to that of CD than to that of AB, it can never have formed a group with AB in opposition to the other manuscripts, and we may, therefore, safely replace Σ_{EF} by Σ. The same argument applies, *mutatis mutandis,* to the second factor likewise, in which Σ_{AB} may equally be replaced by Σ. It follows that where divergence is independent resolution is always legitimate, and, looked at from such a point of view as to make the divergence independent, the variant $AB : CD : EF$, in the particular example selected, becomes $(\Sigma : AB) . (\Sigma : EF)$.

Let us now look at the same readings from the point of view of AB, writing this as Σ. If this group preserves the original reading it is clear that that of CD must have arisen from it, and that of EF by further variation from that of CD. Thus the instance is one of *successive variation.* Here it is evident that the complex variant $\Sigma : CD : EF$ is

the product of the simple variant Σ : CDEF and a further variation of EF from CD. But Σ : CDEF is, of course, equivalent to Σ : AB, and, since variation in EF might have occurred without the previous variation in CDEF, the variant CD : EF is equivalent to Σ : EF. The formal proof would be thus: Σ : CD : EF when factorized becomes

$$(\Sigma_{EF} : CD) . (\Sigma_{CD} : EF).$$

In the second factor, since the readings of AB and CD are closer than those of AB and EF, the latter pair cannot have formed a group apart from CD, and we may, therefore, safely replace Σ_{CD} by Σ. The opposite is the case in the first factor, which cannot, therefore, be simplified so directly. But Σ_{EF} : CD is equivalent to Σ_{EF} : AB, and since the readings of EF and CD are closer than those of EF and AB, we may now replace Σ_{EF} by Σ. Thus $(\Sigma_{EF} : CD) . (\Sigma_{CD} : EF)$ is equivalent to $(\Sigma : AB) . (\Sigma : EF)$.

Since, looked at from the point of view of EF the case stands on the same footing as it does looked at from that of AB, the resolution is of course the same. Consequently, where divergence is successive resolution is likewise always legitimate, and, looked at from such a point of view as to make the divergence successive, the variant AB : CD : EF, in the particular example selected, becomes

$$(\Sigma : AB) . (\Sigma : EF).$$

Thus, as we should have expected, given certain readings showing a completely determinate order of divergence, the resolution of the complex variant will be the same whether we look upon it as independent or divergent. The point of view will be immaterial. In practice then we may adopt the simple rule to put Σ for the intermediate reading, and construct our two factors by opposing Σ with the two extremes. But it must not be supposed from this that, when the point of view is given, it

is indifferent, for the purpose of resolution, whether the divergence is independent or successive. It is evident indeed that the reverse is the case. Let us indicate the point of view by Σ in the variant Σ : CD : EF, then if the divergence is successive the resolution will be (Σ : AB) . (Σ : EF), whereas if the divergence is independent the resolution will be (Σ : CD) . (Σ : EF).

Next, as an instance of *indeterminate divergence*, let us take the readings

Σ *To you I tell* (*xyyx*)
CD *To you I say* (*xyyy*)
EF *To you I sing* (*xyyz*).

Here no two are more closely similar than any other two, so that no order of divergence can be established. Supposing Σ to preserve the original reading there is nothing to show whether those of CD and EF arose from it independently or one through the other, and, if so, which through which. Hence the term indeterminate. Σ : CD : EF when factorized becomes $(\Sigma_{EF} : CD) . (\Sigma_{CD} : EF)$, but, since we know nothing of the affinities of the readings, it is impossible to get rid of the qualified symbols, and no effective resolution can be obtained. Where, therefore, divergence is indeterminate significant variants of complex type are insoluble.

Cases sometimes arise in which, though there seems to be some order of divergence, it is not complete, and may be called *semi-determinate*. As an example, suppose that we take the readings

amorous AB : *loving* CD : *lufand* EF
(*xyxx* AB : *xyxy* CD : *xyyx* EF).

Here there can be no doubt that the readings of CD and EF resemble one another more closely than either of them resembles that of AB, but at the same time it is not possible to say that any one

is intermediate between the other two. If we suppose that of AB to be original, we may reasonably assume that those of CD and EF did not arise from it independently, but one through the other, though there is nothing to show which was the earlier. We may call this *quasi-successive divergence.* If we try to resolve the variant, we shall find that our second factor CD : EF is ambiguous, representing either Σ : CD or Σ : EF (or, in the formal proof, we shall find that neither Σ_{EF} : CD nor Σ_{CD} : EF will simplify). If on the other hand we suppose one of the other readings to be original, say that of EF, then those of AB and CD may have arisen from it independently, or that of AB may have arisen through that of CD. We may call this *quasi-independent divergence*, but it is not distinguishable, so far as resolution is concerned, from the indeterminate type.[1]

Our results may be summarized as follows. Whenever the proposed factors of a complex variant are of type 1, the resolution is always legitimate.[2] When, however, one or both of the proposed factors are of type 2, no resolution is legitimate if the divergence is indeterminate (or only semi-determinate), but if a complete order of similarity between the readings can be established, that is, if the divergence is determined as either successive or independent,

[1] In some particular cases resolution would appear to be possible. Thus when divergence is quasi-successive it would seem legitimate to resolve such a variant as Σ : DE : F as (Σ : DEF) . (Σ : F), and such a variant as Σ : D : EF as (Σ : DEF) . (Σ : D). On the other hand, when divergence is quasi-independent, it would still seem legitimate to resolve Σ : D : EF as (Σ : D) . (Σ : EF), but not Σ : DE : F as either (Σ : DEF) . (Σ : F) or (Σ : DE) . (Σ : F). But in any case these particular varieties, even if theoretically soluble, are hardly worth considering.

[2] Where, however, the divergence is determinate these cases are best treated in the same way as the rest. If the divergence is independent the result will be the same, but it will differ if the divergence is successive.

resolution is always legitimate, though it will differ in the two cases. In practice, whenever divergence is determinate, it is convenient to look at it from such a point of view as to make it independent and to resolve accordingly.

The discussion above has proceeded on the supposition that our complex variants were of type 3. They may, of course, be of that or any higher type. As we proceed the questions of divergence and resolution naturally become more complicated, but as no fresh principles seem to be involved, the matter need not here be pursued, especially as instances where the divergence is completely determinate in variants of higher type than 3 are naturally rare.[1]

Thus we see that not all variants of complex type are immediately soluble, but this results not from anything in the nature of variation, but simply from the fact that divergence is sometimes indeterminate, that is to say, that it does not always appear from the nature of the readings whether, assuming some one to be original, the variations by which the others arose were successive or independent. One or other they must have been, and as soon as we know which, all variants can be resolved.[2]

Wherever divergence is determinate there is a definite *order* among the readings. This may be written X, Y, Z, when the reading Y is intermediate between the other two, and in respect to a particular variant the order of the manuscripts may be deduced

[1] We saw before that to produce completely determinate divergence in a variant of type 3 it would appear that the reading must contain two independently variable elements. It would, therefore, seem that a type-4 variant would need three such elements, and so on. But there may be complications.

[2] It is well to insist upon this, since it is sometimes tempting to suppose that insoluble variants point to independent variation. This is not so. It is only lack of evidence, never the nature of derivation, that makes variants insoluble.

from that of their readings, and written X, Y, Z. This order is, of course, the same as Z, Y, X. From the point of view of Y the divergence is independent, from those of X and Z successive. When the point of view is determined by the establishment of originality *direction* is introduced. Thus, if the reading of X is original the ordered series X, Y, Z will become the directed series $X > Y > Z$, if that of Y then $X < Y > Z$. Here it must be clearly understood that the direction applies primarily to the readings and only indirectly to the manuscripts. If we write $X > Y > Z$, what we mean is that the reading of Z is derived from that preserved in Y, and the reading of Y from that preserved in X, not that Y is itself derived from X, or Z from Y. If the descent of the manuscripts themselves were intended we might write $X \rightarrow Y \rightarrow Z$.

Order and direction are notions that apply to individual readings. It is clear that in some sense they must also apply to several or all the readings of the different manuscripts and thus to the manuscripts themselves. However, the generalization of these notions presents very considerable difficulties, and, since the knowledge that could be gained from it does not differ in nature or extent from that to be derived from a collection of simple variants, there would be no object in trying to substitute it for the far more convenient method of the compounded variational formula, and the possibilities need not be further considered.

At the end of this lengthy discussion it may seem like a bad joke if I add that the question of resolution is, after all, of secondary importance. Yet this is, in a manner, true. It seems probable that in most cases the natural variants of type 2 will prove sufficient to establish the manuscript relation, and that no material accession of evidence will result from the resolution of complex types. On the other

hand it may well happen, in the case of some texts preserved in a large number of manuscripts, that variation has advanced to such a point, that practically all variants are of complex type. And, in any case, for the complete elucidation of the problem resolution will always be necessary. There is a double reason for this. In the first place, we require to make sure that the complex variants, especially if relatively numerous, do not conceal any anomalous groupings, such as would either invalidate our inferences as to relationship, or reveal the presence of conflation, as I shall explain later. And further, resolution is needed to enable us to ascertain the total number of cases in which a manuscript is necessarily in error, and so to place it among its fellows in order of merit—another question that will engage our attention before we have done with the calculus.

Method and Limitations of the Calculus

When the variants have all been recorded, and, wherever possible, resolved into their simple factors, the next step is to sort them into similar classes, each comprising a single form alone. If all is well, these classes will be constant, that is to say that, out of all possible forms of type-2 variants,[1] certain ones will predominate to the practical exclusion of others: furthermore the predominant classes will be mutually consistent.

If all is not well, that is, if the grouping is throughout random or if inconsistent forms are of frequent occurrence, the relationship of the manuscripts cannot be accounted for on the hypothesis of simple transcription; some sort of conflation has somewhere to be assumed (i. e. we must suppose that at some

[1] All forms of type 1 should occur, otherwise the collection of manuscripts examined will not be purely collateral.

point the genetic relation has been, not one-one or one-many, but many-one). This is a matter that lies beyond the scope of the calculus and cannot be properly discussed in these pages: since, however, the application of the calculus to the problem raises some rather interesting speculations, I have ventured to touch on the matter in a final note.[1]

The extent to which inconsistent grouping may be expected to occur when there is no conflation, will depend upon the degree of collation, as previously explained. The minuter the collation the greater will be the number of abnormal variants, not only absolutely but relatively. When collation is confined to variants of real importance, every anomalous grouping should be capable of definite explanation.

We next turn to those variants which, owing to indeterminate divergence, were not immediately soluble, and resolve them in accordance with the groups already established, that is, taking care that no groups shall be produced inconsistent with those that already occur.[2] Our collection of normal variants is now complete,[3] and they belong, so far as they are significant, to type 2 alone.

We next proceed to compound the variants in the manner already explained.[4] Thus, if the only constant groupings are $\Sigma : AB$, $ABC : DEF$, $\Sigma : EF$, we write $\Sigma : C[D(EF)]$; if $\Sigma : AB$, $\Sigma : CD$, $\Sigma : EF$, then $\Sigma : (CD)(EF)$. In regard to this one point

[1] See Note B.

[2] Where not only is the divergence indeterminate, but the constant grouping consistent with quasi-independent derivation, the resolution may still be ambiguous. For example, if the grouping is $\Sigma : (CD)(EF)$, then in resolving the variant $AB : CD : EF$ there is nothing to show which pair of factors to select. In such a case all three factors should be given, viz. $(\Sigma : AB).(\Sigma : CD).(\Sigma : EF)$. This may actually represent the facts, and if not it at least weights the different groups equally.

[3] Anomalous groupings can only be dealt with when the family relation has been established (see below, p. 51).

[4] See p. 23.

needs to be made clear. It must be remembered that we are dealing merely with variational groups, and we must never introduce more brackets than are needed to express these groups. It need hardly be pointed out that the grouping $\Sigma : C[D(EF)]$ could equally well be expressed as $\Sigma : D[C(AB)]$; but it may perhaps be well to observe that, although it might be tempting to write the formula in the more definite form $AB\{C[D(EF)]\}$, this would be illegitimate, since it would suggest that C was more closely related to one of the alternative groups AB and DEF than with the other (though leaving ambiguous which), and this there is no reason to suppose. It is not without design that one ':' has been retained in these formulas instead of being replaced by a bracket, since it serves to indicate a real limitation.

We now desire to take the all-important step from variational to genetic groups, and, from the observed affinities of extant manuscripts, to infer the ancestral family. Suppose, for a moment, that the only constant significant groupings are

$$\Sigma : AB,\ ABC : DEF,\ \text{and}\ \Sigma : EF,$$

and the compounded formula, therefore,

$$\Sigma : C[D(EF)].$$

This is, of course, the grouping that will result from a family relation defined by such a parallel formula as $(x)A'A\{B[C(D\overline{EF})]\}$. This is important, no doubt, but insufficient for our purpose: we require to know, not merely that a particular relation will account for the grouping, but that it alone will do so. Is the family just defined unique in this respect?[1] Since in it CDEF is a genetic group,

[1] Of course, since the groupings are equally expressed by $\Sigma : D[C(AB)]$, and this may be written $[(AB)C]D : \Sigma$, it follows that they are equally explained by the family $(x)A'\{[(\overline{ABC})D]E\}F$. Thus the family considered above is clearly not unique.

let us replace it by xA'CDEF, say ϕ. We are then left with the three manuscripts, A, B, and ϕ, and are therefore confronted with the familiar ambiguity. Any one of the three possible assumptions,

$$(x)A\text{'A(B}\phi), \quad (x)A\text{'(A)(B)}(\phi), \text{ and } (x)A\text{'(AB)}\phi,$$

will equally well account for the groupings we have supposed. The first of them is, of course, the family with which we started above; the others will be defined by the formulas

$$(x)A\text{'}\{\text{A}\}\{\text{B}\}\{\text{C [D(EF)]}\} \text{ and}$$
$$(x)A\text{'}\{\text{AB}\}\{\text{C [D(EF)]}\}$$

respectively. Nor is this the end of our uncertainty. For in the last of the new families there exist both xA'AB, say α, and xA'DEF, say δ, and consequently the ambiguity repeats itself with the manuscripts α, C, and δ. The alternatives now are:

$(x)A\text{'}\alpha(\text{C}\delta)$ = as above,
$(x)A\text{'}(\alpha)(\text{C})(\delta) = (x)A\text{'}[\text{AB}][\text{C}][\text{D(EF)}]$, and
$(x)A\text{'}(\alpha\text{C})\delta = (x)A\text{'}[(\text{AB})\text{C}][\text{D(EF)}]$.

Again, in the last of these families, the existence of xA'ABC, say β, and of xA'EF, say ϵ, lead to the alternatives:

$(x)A\text{'}\beta(\text{D}\epsilon)$ = as above,
$(x)A\text{'}(\beta)(\text{D})(\epsilon) = (x)A\text{'}[(\text{AB})\text{C}][\text{D}][\text{EF}]$, and
$(x)A\text{'}(\beta\text{D})\epsilon = (x)A\text{'}\{[(\text{AB})\text{C}]\,\text{D}\}\{\text{EF}\}$.

Finally, in the last of these families, the existence of xA'ABCD, say γ, leads to the alternatives:

$(x)A\text{'}\gamma(\text{EF})$ = as above,
$(x)A\text{'}(\gamma)(\text{E})(\text{F}) = (x)A\text{'}\{[(\text{AB})\text{C}]\,\text{D}\}\{\text{E}\}\{\text{F}\}$, and
$(x)A\text{'}(\gamma\text{E})\text{F} = (x)A\text{'}\{[(\overline{\text{AB}}\text{C})\text{D}]\text{E}\}\text{F}$.

This last will be recognized as the exact reverse of the family with which we began. Thus so long as

we depend upon variational groups alone, we get, as indeed we should expect, a continuous series of possible schemes of relationships connecting our extreme cases, the common condition being that there should never be more than three manuscripts, and these not all inferential, independently derived from the archetype.

As another example let us consider the family $(x)A'(AB)(CD)(EF)$. The normal type-2 variants in this case will be $\Sigma : AB$, $\Sigma : CD$, and $\Sigma : EF$. Of these $\Sigma : CD$ is inconsistent with $ABC : DEF$ and the family, therefore, found no place in our former series. Here, since $xA'AB$, $xA'CD$, and $xA'EF$ all exist, we are at once confronted with the familiar ambiguity, and the same variational formula, viz. $\Sigma : (CD)(EF)$, will result from any of the relations defined respectively by

$(x)A'(AB)(CD)(EF)$, as above,
$(x)A'[AB][(CD)(EF)]$, and
$(x)A'[(AB)(CD)][EF]$.

It will be realized how severe is the limitation which the ambiguity of three texts imposes on the application of the calculus. To solve the difficulty it is necessary to introduce the notion of direction of variation, which, as already said, itself lies outside the calculus, since it involves judgements concerning the originality of readings, which are incapable of logical proof.

The question of originality bears directly on the problem of the three texts. The ambiguity arose because, since variation might arise in any manuscript, the type-1 variants, which were the only simple ones possible, could always be accounted for on the supposition of individual variation; no significant grouping could ever be established. But if it can be shown that one manuscript preserves an original reading where the others agree in the same

unoriginal reading, then it will follow that these form a genetic pair and are derived in common from another manuscript in which the variation arose. The manner in which the originality or unoriginality of readings can be rendered probable is a subject that necessarily occupies a prominent place in textual criticism, but which I do not propose to deal with here. There are, however, one or two formal points connected with it that deserve a word in passing. Since nothing more than probability can be established touching the originality of any individual reading, no inference as to genetic grouping should be accepted which is not based on an examination sufficiently critical and extensive to eliminate any sensible risk of deception, on the one hand by chance coincidence in error, and on the other by the ingenuity of editorial emendation, two forms of what we have called convergent variation. It must also be borne in mind that the fact of some divergent readings in a manuscript being original affords no ground for supposing that all are: some are certain to be individual errors. Indeed, it may easily happen that the only manuscript that preserves original readings in its type-1 variants is nevertheless the most corrupt of all. Of course, of any three manuscripts, not more than one can be original in its simple variants.

Let us now see how the introduction of the notion of originality affects the ambiguity which we found inherent in our compounded formula. We perceived that in Σ: C[D(EF)] there was uncertainty as to the affinity of C, and a study of the transmutations of our hypothetical families will show that the formula covers the three possibilities, {AB}{C[D(EF)]}, [AB][C][D(EF)], and [(AB)C][D(EF)]. If, therefore, we can ascertain that the readings of CDEF are sometimes unoriginal, we shall have proved that it is the first of these alternatives that represents

the facts. We can, moreover, extend our analysis further. Putting γ for xA'CDEF, if we find that, in simple variants, γ is sometimes original, then AB is a genetic group; if A or B is sometimes original, then Bγ or Aγ respectively are genetic groups; if neither A, B, nor γ is ever alone original in simple variants, then all three are independently derived. It should also be observed that, if EF (or any other pair) is a genetic group, the order within it must be permanently indifferent, since sometimes one and sometimes the other manuscript will have varied the reading of the parent.[1]

We may now return to the possibility of inferring the family relation from the observed grouping, and consider afresh the hypothetical arrangements capable of producing given groups. And now, of course, the solution of the three-text ambiguity, by the introduction of the notion of originality, enables us to avoid the kaleidoscopic transformations in which we previously found ourselves involved. But it may be useful to consider exactly how much direction we need to establish in order to define a particular family. If the choice lay between the two extreme cases of our former series, namely the

families $(x)A'\mathrm{A}\{\mathrm{B}[\mathrm{C}(\mathrm{D}\overline{\mathrm{EF}})]\}$

and $(x)A'\{[(\overline{\mathrm{AB}}\mathrm{C})\mathrm{D}]\mathrm{E}\}\mathrm{F}$,

it would be sufficient to satisfy ourselves that the significant readings of any one manuscript were consistently more or consistently less original than those of some other later in the series, in order to establish the first or second respectively. But this

[1] To this extent the order in a genetic formula is indeterminate. Within any bracket the order of the terms of the same order is indifferent. Thus, in $(x)A'\mathrm{A}\{\mathrm{B}[\mathrm{C}(\mathrm{D}\overline{\mathrm{EF}})]\}$, it is indifferent whether we write $\overline{\mathrm{EF}}$ or $\overline{\mathrm{FE}}$, $(\mathrm{D}\overline{\mathrm{EF}})$ or $(\overline{\mathrm{EF}}\mathrm{D})$, and so on. In the absence of any other determining factor, such as date, we naturally give such an order to our manuscripts as is likely to lead to some symmetry in the formula.

would not suffice to exclude intermediate cases. It will be seen that, given the family

$$(x)A'\mathrm{A}\{\mathrm{B}[\mathrm{C}(\mathrm{D}\overline{\mathrm{EF}})]\},$$

in variants of the form $\Sigma : \mathrm{AB}$, the reading of AB must be original, while, in variants of the form $\Sigma : \mathrm{A}$, originality must lie sometimes with A and sometimes with Σ. It follows that, if a reading found in AB alone is always, and a reading found in A alone is sometimes, original, then, from the variational grouping

$$\Sigma : \mathrm{C}[\mathrm{D}(\mathrm{EF})],$$

we are able definitely to infer the above family. And more generally we may say that, if any collection of collaterals can be arranged in such an order that all type-2 variants divide the collection into groups one of which is wholly anterior to the other, and if, of such complementary groups, the earlier is always the more original, and further if, of the two manuscripts constituting the earliest group, one is sometimes original in its type-1 variants, then it follows that the family is completely successive, as in the case of $(x)A'\mathrm{A}\{\mathrm{B}[\mathrm{C}(\mathrm{D}\overline{\mathrm{EF}})]\}$.

On the other hand, given once more the grouping $\Sigma : \mathrm{C}[\mathrm{D}(\mathrm{EF})]$, and supposing the groups ABCD and CDEF to be consistently more original than their complements, while of ABC and DEF sometimes one and sometimes the other is original (and the inconsistent group CD duly absent), then it will follow that the family is $(x)A'[(\mathrm{AB})\mathrm{C}][\mathrm{D}(\mathrm{EF})]$. In practice, of course, we try to establish as many groups of greater originality as possible, in order to render the relation of the manuscripts certain.[1]

[1] Groups of greater originality are, of course, the complements of the groups of common error familiar in textual criticism, if by 'error' is meant a departure from the archetype, and not, as is often the case, merely an incorrect reading. Many so-called errors are original readings of the archetype, and their plausible alternatives unoriginal emendations.

What has been said above is only valid provided the whole collection is, as we have assumed, collateral throughout. In practice, when we have reached this point, it will be well to reconsider this question, since it is just with respect to a single independent manuscript, such as A in the assumed family $(x)A'A\{B[C(D\overline{EF})]\}$, that error is likely to occur. Suppose we have arrived at the existence of $xA'BCDEF$, say β, then there are three possibilities as to the relation of A and β: namely, A and β may be identical ($A = \beta$), or β may be derived from A ($A \rightarrow \beta$), or A and β may be collateral ($A'A\beta$). In each case there will be distinguishing features, by which our assumption may be tested: if $A = \beta$, then $\Sigma : A$ cannot occur; if $A \rightarrow \beta$, then in $\Sigma : A$ the reading of A must be original; if $A'A\beta$, then in $\Sigma : A$ sometimes A and sometimes Σ will preserve the original reading.

Having now determined the family relation of the manuscripts, we must return to the consideration of any anomalous variants there may be. If they prove easy of explanation, as due to convergent error or emendation, all is well, and we may convert them into the normal groupings of which they appear to be concealed examples. If, on the other hand, they prove intractable, and a reconsideration of the manuscript relation suggests no solution, we shall have to suppose that some conflation has occurred.

Given the relationship of the manuscripts, it is easy to specify which groups, either of extant or inferential manuscripts, are necessarily original in their variant readings, which necessarily unoriginal, and which uncertain. Readings supported by necessarily original groups are immediately available for editorial purposes; it is those supported by groups of uncertain value that give trouble. They have, of course, to be treated on their merits, but the

calculus can, nevertheless, help us by establishing an order of merit among the manuscripts. To do this we take our completed list of simple readings and analyse it to ascertain the number of necessarily original (or unoriginal) readings in each manuscript, extant or inferential. The evidential value of each manuscript, or group of manuscripts, thus receives numerical expression.

To illustrate this, take once more the family $(x)A'A\{B[C(D\overline{EF})]\}$. In this any group of two or more manuscripts including A must preserve the reading of the archetype; while, if A is unavailable, any group of two or more including B must preserve the reading of xA'BCDEF, which is as near to the original as we can get. The reading of the complement of such a group is, of course, necessarily unoriginal. The errors, say, of D are the readings of all the unoriginal groups in which D appears; namely, the specified groups in $\Sigma:D$, $\Sigma:DEF$, $\Sigma:CDEF$. The reading of A alone (i.e. in $\Sigma:A$) is always uncertain, and so is that of B in $\Sigma_A:B$, whether A is really absent or whether the formula arises through resolution. This is why we must be careful not to reject the divergent reading in $\Sigma_A:B$ on the ground that in $\Sigma:B$ it is necessarily unoriginal.[1]

[1] Cf. p. 28[2]. The numerical expression of value must not be taken as more than closely approximate. This is due in part to the fact that, if we adopt the method recommended above (p. 28) of removing qualified symbols wherever possible, a manuscript by making an extensive omission (counting as a single variant) escapes being charged with a number of errors it would probably have otherwise made. (This difficulty could, of course, be met by making an allowance of errors for the omission.) It is also partly due to the fact that we decided (p. 44[2]) that where an error could not be definitely located it was to be assigned to all possible groups. A further limitation arises from the fact that a manuscript may be very independent in minor matters, and therefore appear of small evidential value, and yet be a trustworthy guide in more important readings. Of course, it may happen that the

The relative authority of A and xA^{ϵ}BCDEF can never be numerically determined, but, in the family in question, values can be given to all the other manuscripts, and these values the occasional absence of A does not disturb.[1] Consequently, when A is unavailable we have some *a priori* guide to our choice between the reading of B and of xA^{ϵ}CDEF. It must not be supposed that this is of great importance in individual cases. To know, say, that, in the variant Σ_A : B, the chances are two to one in favour of Σ,[2] cannot seriously affect our judgement of the inherent merits of the readings. It affords, however, a useful check over a sufficient field. Unless, in a large number of cases, our choice between Σ and B agrees approximately with their numerical values, we may suspect that we have been somehow biassed in favour of one or other, most likely by the ingenious editorial activities of a scribe.

This completes the present sketch of the Calculus of Variants. Its scope is admittedly restricted, since, without the notion of originality, which has to be imported from outside, it can lead to no definite results. Moreover, the introduction of this notion of originality is attended with very great logical difficulties. For it is not sufficient, in order to attain definiteness, to establish originality in a

most important, because genetically independent, of the extant manuscripts (e. g. A in the family considered above) may at the same time be the most corrupt.

[1] Hence the importance of an occasionally available manuscript.

[2] The chances, in such a case, will generally be in favour of Σ. This is due to the fact that xA^{ϵ}CDEF may very likely be an immediate transcript of xA^{ϵ}BCDEF, whereas between the latter and the extant B there probably intervene a number of potential manuscripts. But a series of transcriptions is always likely to introduce more variation than a single transcription, and Σ has therefore a better chance at being original than B. This is the grain of truth that lies behind the fallacy of constant variation (see Note C). But, of course, B may be as old as xA^{ϵ}CDEF.

few cases sporadically, we need to do so in every instance of certain given classes.[1] But if in every case we can determine by inspection which reading is original, the relation of the manuscripts ceases to be of any interest whatever. This means that the reconstruction of the text of an archetype can never be certain, but is always at best a matter of greater or less probability. This may be, and I suppose universally is, admitted; but the admission does not appear to afford any ground for rejecting the help of a rigorous logic so far as it can be made available. That it is capable of lending material assistance is suggested by the singular errors into which critics sometimes fall.[2]

[1] The fundamental case is, that, in order to establish the independent derivation of three manuscripts, we need to show that every reading supported by two manuscripts is original. What, of course, we do in practice is to assume that, if, in all cases in which originality can be reasonably guessed at, it lies with a particular group, then it always lies with that group. How far this procedure is likely to lead to correct results depends, naturally, on the proportion of instances in which originality can be presumed.

[2] A few observations on this head will be found in Note C. I may add that nothing has more forcibly brought home to me the ease with which one may trip in pursuing this subject than the many fallacies I have detected in successive drafts of the present essay.

NOTE A—*On Collateral Groups*

Throughout the discussion in the text, whenever it was necessary to take a specific collection of manuscripts for the purpose of illustration (viz. A, B, C, D, E, and F), it was assumed that this formed a collateral group, that is to say, that no manuscript of the collection was an ancestor (or descendant) of any other. The operation of eliminating descendants from a collection of extant manuscripts is not one that can be performed by the use of the calculus alone. Nevertheless, it may be well, while not attempting to deal fully with the matter, at least to inquire how far the calculus is able to help. In practice, the fact of one extant manuscript being a copy, mediate or immediate, of some other, is often apparent from evidence of which the calculus is unable to take account, and the elimination of derivatives is perhaps not generally a very difficult task.

The test, a useful if incomplete one, that the calculus affords for the detection of derivatives is the absence of type-1 variants. Suppose, for a moment, that our six manuscripts formed an ancestral group or series, with descent from A to F_6 in alphabetical order, then the only type-1 variants that could occur would be $\Sigma : A$, due to variation in the writing of B, and $\Sigma : F$, due to variation in the writing of F. Hence it seems likely that the presence of descendants will reveal themselves by the absence of certain forms of type-1 variants. If, for instance, some one of our manuscripts, say F, is derived from some other, say E, we shall get variants of the form $\Sigma : EF$, due to variation in E, and variants of the form $\Sigma : F$, due to variation in F, but (resolution apart) we shall get no variants of the form $\Sigma : E$. However, the same result as regards variants would follow if F and xA^{c}ABCD were independently derived from E. If we are able to establish direction of variation, the two cases will, of course, be clearly distinguished, since, in the latter case but not the former, the reading of E, whatever the form of the variant, must be original. Again, if only type-1 variants occur, and one particular form, say $\Sigma : A$, is absent, it follows,

either that all the other manuscripts are independently derived from A, or else that A is itself derived from one other manuscript and is the parent of all the rest. In the former, but not the latter, case the reading of A would be necessarily original. It follows, of course, from the postulate of universal variation that the absence of one or more forms of type-1 variants is inconsistent with the collection being purely collateral, though the determination of direction of variation is needed to show in what particular manner it is ancestral. The converse, however, is not true: the presence of all forms of type-1 variants, irrespective of higher types, does not prove the collection to be throughout collateral. For if A, for instance, is the parent of the manuscript from which all the rest are derived (i.e. $A \rightarrow xA'BCDEF$), then Σ: A will occur through variation in that common ancestor. The elucidation of this case in the light of direction of variation has already been considered in the text (p. 51: the reading of A is, of course, necessarily original).

NOTE B—*On Conflation*

By *conflation* I understand the appearance in a manuscript of readings which are neither derived from the archetype (by continuous descent) nor are original variants of its own or any of its ancestors, but have been imported from some other line of descent.[1]

There are two ways in which conflation may arise, and they produce somewhat different results: they may be called correctional and editorial respectively.

Correctional conflation would appear to be much the more common. It happens through a manuscript being collated with, and absorbing readings from, some other manuscript. Sometimes the foreign readings are actually inserted in the text over erasure of the original readings, in which case the manuscript itself becomes conflated. More often, perhaps, the readings are written in the margin or between the lines, and copied into the text in the course of transcription, in which case it is the derivatives that show conflation. If this process of collation and 'correction' has been sufficiently careful and minute the result may be practically to transfer

[1] Some critics term this 'mixture' or 'contamination', reserving 'conflation' for the particular case in which the readings of two lines of descent appear side by side in the text.

the manuscript or its descendants from one line of descent to another, and the calculus may give no warning of what has occurred. This, however, would be an extreme, and in practice a very unlikely, case. What usually happens is that collation and 'correction' are confined to some of the more striking variants. This will show itself on analysis either by the sporadic appearance of anomalous groupings, or by those involving the more important variants consistently pointing in one direction, and those involving the minor variants consistently in another. In either case, the calculus may be expected to throw some light on the problem. It may be added that, where conflation is suspected, the value of variants as an indication of ancestry is in inverse proportion to their intrinsic importance. To the herd of dull commonplace readings we must look for the genetic source of the text, to the more interesting and striking for the source of the contamination. Nothing can be more misleading than to seek to 'place' a manuscript on the evidence of a few 'test' readings.

Editorial conflation occurs when a scribe while writing a manuscript has several others of the same work open before him. If he is thorough enough, and digests the readings of his exemplars in every phrase as he writes it, adopting at the same time his own spelling and grammatical usage, he may not only produce a text upon whose affinities the calculus can throw no light, but one in connexion with which the question of ancestry has no significance. He would be acting, in fact, like many modern editors, whose texts, however readable, afford no guide to what the author may have written. At the other extreme, he may copy consistently from one exemplar, only introducing readings from some other when he is dissatisfied with the text presented. In that case his work will be indistinguishable from that of the correctional conflationist. But what is most characteristic of editorial conflation is a tendency to jump from one exemplar to another. The scribe copies closely from one of the manuscripts before him, till a sudden dissatisfaction with some reading, or it may be mere chance, sends him off to another, which he then follows for a while, before either passing on to a third or perhaps returning to the first. The result, in a broken and confused manner, recalls those composite manuscripts of which different portions belong to different lines of descent. It may to some extent be possible to follow the scribe in his vagaries

and disentangle the nature of his several exemplars, but in this task we can hardly look for assistance from the calculus.

NOTE C—*On Some Common Errors*

One main object of the calculus is to define unambiguously the notions required in textual criticism, and by the more rigorous methods of symbolic treatment to obviate some of the errors into which critics appear to have fallen. It would be easy to cite examples of the lax use of technical terms and of statements to which no precise meaning can be attached. Only too often one is driven to conclude a writer's meaning by a process of excluding those which, on an assumption of ordinary intelligence, he may be supposed not to have intended. And it sometimes happens that when some vaguely plausible assertion has by sufficient assumptions been made definite, it is seen to be no longer plausible at all. Much needs only to be defined to be rejected. But rather than pursue so obvious a complaint I will mention briefly two matters in which confused ideas appear to have led to unsatisfactory results.

The first is the danger of incomplete collations. Manuscripts are, as a rule, collated separately with some standard text, often in circumstances that do not permit of subsequent verification. Such collations are useless for the purpose of establishing textual relationship, since an incomplete collation not merely fails to give full information, but gives information that is actually false. Take once more for illustration the family $(x)A^{\epsilon}A\{B[C(D\overline{EF})]\}$, and let A be the manuscript used as the basis of collation, with which the others are compared. Now, suppose that while C and D have been fully collated, the examination of E and F has been so defective that some variants common to both have been overlooked. In so far as these are variants of EF only, or even of DEF, it will not matter, beyond affecting the numerical expression of the value of the manuscripts. Should any, however, be variants of CDEF, the oversight will have the result of substituting the formula $\Sigma : CD$ for the correct $\Sigma : AB$ $(= \Sigma : CDEF)$. Thus a purely fictitious group CD will be created, which may lead to the assumption of a family $(x)A^{\epsilon}(AB)(CD)(EF)$; while the genuine variants $ABC : DEF$ will appear anomalous and very likely be put

down to conflation. Or again, suppose that variants of D have been overlooked. Then, so far as these are variants of D only or of DEF, it will not matter. But should they be variants of CDEF or of BCDEF, then the fictitious groups ABD or AD will be created, with equally misleading results. Such errors also have an unfortunate reaction, for a critic, observing that some manuscript, of an admittedly inferior genetic group, at times (apparently) preserves an original reading, is led to attach a fictitious importance to its readings when other superior manuscripts are defective or unsatisfactory. The only safe method of collation is, first to compare each manuscript severally with the standard, and then to check the whole list of variants (in whatever manuscript) with each manuscript in turn. Only so can one make reasonably certain that, if a variant reading is quoted at all, it shall be quoted for all manuscripts in which it occurs. The total failure to record a variant reading is comparatively harmless; to specify some of the manuscripts containing it and not others can only lead to disaster.

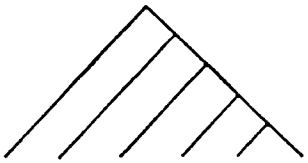
The other point I wish to mention is the error of supposing that every transcription introduces approximately the same number of variants.[1] This, of course, is never actually asserted, nor perhaps consciously entertained, but it would appear to be sometimes tacitly assumed. For instance, we may find it stated that of the three manuscripts X, Y, and Z, the first two are the more closely related, and in explanation we are told that there are more readings common to XY than to any other pair. If this is all that is *meant* by the statement, it is a misuse of the term 'related'. At times, however, it seems to be supposed that the fact really implies a genetic relation. The argument, made explicit, would apparently run somewhat as follows: Of a total of 100 variants, 50 are of the form XY : Z, 25 of the form X : YZ, and 25 of the form XZ : Y; therefore XY have a common source χ, and χZ a common source ζ, each of the transcriptions $\zeta \rightarrow \chi$, $\zeta \rightarrow$ Z, $\chi \rightarrow$ X, $\chi \rightarrow$ Y introducing the same number of variants, namely 25. Of course, the implicit assumption is unfounded, and the whole process illegitimate. The explanation may equally well be that all three manuscripts are independently derived from a common source, and that the scribe of Z was twice as careless as the scribes of X and Y.

[1] The fallacy of constant variation mentioned above, p. 9 note.

DIAGRAMS OF TYPICAL FAMILIES OF THE SIX MANUSCRIPTS A, B, C, D, E, F, TOGETHER WITH THE FORMULAS DEFINING THEM

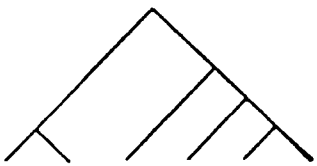
The examples have been selected mainly to illustrate the discussion on pp. 45–7.

$(x)A'A\{B[C(D\overline{EF})]\}$

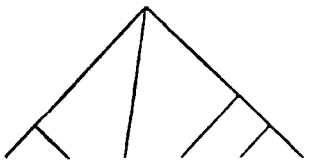


$(x)A'\{A\}\{B\}\{C[D(EF)]\}$

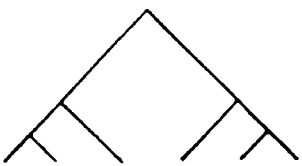
$(x)A'\{AB\}\{C[D(EF)]\}$



$(x)A'[AB][C][D(EF)]$

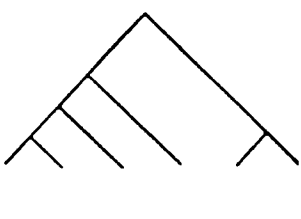


$(x)A'[(AB)C][D(EF)]$

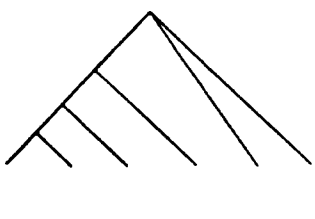


$(x)A'[(AB)C][D][EF]$

$(x)A'\{[(AB)C]D\}\{EF\}$

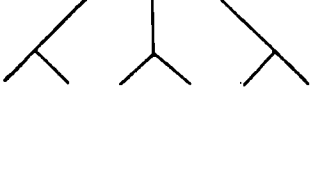


$(x)A'\{[(AB)C]D\}\{E\}\{F\}$

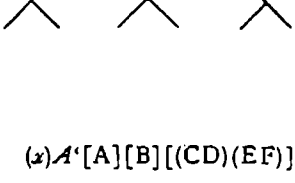


$(x)A'\{[(\overline{A}BC)D]E\}F$

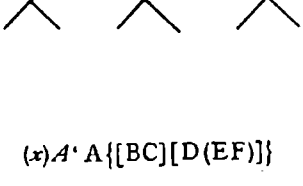
$(x)A'(AB)(CD)(EF)$



$(x)A'[AB][(CD)(EF)]$



$(x)A'[(AB)(CD)][EF]$



$(x)A'[A][B][(CD)(EF)]$

$(x)A'A\{[BC][D(EF)]\}$

INDEX

This is mainly a list of terms, and the references are mostly to definitions.